JOURNAL OF RESEARCH of the National Bureau of Standards
Vol. 85, No. 5, September-October 1980

A Univariate Extension of Jensen's Inequality

Clifford H. Spiegelman*

National Bureau of Standards, Washington, D.C. 20234

May 23, 1980

The result in this paper explains some of the qualitative nature of Jensen's inequality. It is shown that the more disperse the distribution of a random variable is, the smaller is the expectation of any concave function of it. This result can be used to show the inadequacy of some current methods of reporting environmental data by using geometric means, and it extends the result of I. Billick, D. Shier, and C. H. Spiegelman, where symmetry of the error in environmental measurements is assumed.

Key words: Blood lead levels; concave; Jensen's inequality; geometric mean.

The result in this paper explains some of the qualitative nature of Jensen's inequality. It is shown that the more disperse the distribution of a random variable is (as described by conditions *C.* below), the smaller is the expectation of any concave function of it. This result can be used to show the inadequacy of some current methods of reporting environmental data [1],[1] and extends the result of I. Billick, D. Shier, and C. H. Spiegelman [2].

The result of this paper is:

THEOREM: *Let* F *and* G *be two distribution functions such that*

$$\text{C.}\qquad \begin{aligned} &\int X dF(X) = \int X dG(X) = \mu \text{ and} \\ &F(X) \geqslant G(X) \text{ for } X \leqslant \mu \\ &F(X) \leqslant G(X) \text{ for } X \geqslant \mu \end{aligned}$$

Then for any concave function $L(X)$ defined on an open interval of the real line containing the support of F and G,

$$\int L(X)dF(X) \leqslant \int L(X)dG(X).$$

It should be noted that if $L(X)$ is a concave function of X then so is $L(X-\mu)$, for any constant μ. Therefore the assumption of equal means is equivalent to the assumption of both means being zero. In addition if $G(X)$ gives all its mass to one point the inequality above reduces to Jensen's inequality.

Before presenting the proof of the theorem it will be useful to give an application.

APPLICATION:

$$Y_1 = X + V_1(X)W$$
$$Y_2 = X + V_2(X)W$$

where X and W are independent random variables, such that $EW = 0$, $EW^2 = 1$, and that Y_1 and Y_2 are positive a.s.. Also suppose that the standard deviations V_1 and V_2 satisfy $V_2(X) \geqslant V_1(X)$ a.s.. Then the theorem may be applied to show that $EL(Y_2) \leqslant EL(Y_1)$ for all concave functions L. When $L(X) = X^{1/n}$ this example shows that decreasing observed values of the geometric mean $(EX^{1/n})^n$, (of blood lead levels) may only mean that the observations have become more noisy.

PROOF: *Without loss of generality we take* $\mu = 0$.

Notice from the hypothesis of the theorem that

$$X \geqslant F^{-1}(G(X)) \quad \text{for } X < 0$$
$$X \leqslant F^{-1}(G(X)) \quad \text{for } X > 0 \quad \text{where } F^{-1}(t) = \inf\{X \mid F(X) \geqslant t\}.$$

Then if $\sigma(X) = F^{-1}(G(X))/X$, $(X \neq 0)$, it follows that $\sigma(X) \geqslant 1$.

Next consider the integral

$$\int L(\sigma(X)X)dG(X) = \int L(X)dF(X) \tag{1}$$

The proof of the theorem proceeds through three steps. However the main idea of the proof can be obtained from the first step. In this step some regularity conditions are assumed to hold. In steps 2 and 3 these regularity conditions are removed.

Step 1.

Assume: 1. The second derivative $L''(X)$ exists a.s. with respect to both distribution functions F and G.

2. The distribution functions F and G are continuous monotonically increasing and have compact support, (i.e. there exists an interval I such that $\int_I dF = \int_I dG = 1$).

Define $h(\sigma(X)) = \int L(\sigma(X)X)dG(X)$. It will be shown that $h(1-\alpha+\alpha\sigma(X))$ is a decreasing function of α for $0 \leqslant \alpha \leqslant 1$.

Consider
$$\partial h([1-\alpha+\alpha\sigma(X)])/\partial\alpha$$
$$= \int L'([1-\alpha+\alpha\sigma(X)]X)X(\sigma(X)-1)dG(X)\,, \; 0 \leqslant \alpha \leqslant 1. \tag{2}$$

Since $(\sigma(X)-1) \geqslant 0$ and, from the concavity of L, $L'([1-\alpha+\alpha\sigma(X)]X)$ is a decreasing function of X, it follows from Kimbal's inequality (which states that the covariance between an increasing and decreasing function of X is non-positive) that

$$\int L'([(1-\alpha)+\alpha\sigma(X)]X)X(\sigma(X)-1)dG(X)$$
$$\leqslant [\int(\sigma(X)-1)dG(X)]^{-1}\int X(\sigma(X)-1)dG(X)$$
$$\int L'([(1-\alpha)+\alpha\sigma(X)])(\sigma(X)-1)dG(X).$$

If it can be shown that the integral term $\int X(\sigma(X)-1)dG(X) = 0$, then the first derivative will have been shown to be everywhere nonpositive. Thus, when this equality is verified the proof of this step will be complete.

It remains only to note that

$$\int X\sigma(X)dG(X) = \int XdF(X).$$

In the next step the additional assumption that L is smooth is dropped.

Step 2. Assumption (2) remains and assumption (1) is dropped.

Since L is concave and hence continuous it may be approximated arbitrarily closely by an interpolating linear spline $S(X)$, which is differentiable a.s. with respect to both F and G. This spline $S(X)$ is equal to $L(X)$ on at least a finite set of points.

Since the interpolating spline satisfies the assumptions of step 1 by taking limits the results hold for the assumptions of this step.

The next step is the last.

Step 3. No additional assumptions to the theorem are required.

Case 1: $\int |L(X)|\, dG(X) = \infty$.

From the concavity of L, $\int L(X)dG(X) = -\infty$.
Since $\int L(X)dG(X) = \int_{-\infty}^{0} L(X)dG(X) + \int_{0}^{\infty} L(X)dG(X)$ at least one of these two integrals is infinite. Take $\int_{-\infty}^{0} L(X)dG(X) = -\infty$. On the region A where L is increasing we have $\int_A L(X)dF(X) \leqslant \int_A L(X)dG(X) = -\infty$ and the conclusion of the theorem holds.

Case 2. $\int |L(X)| dG(X) < \infty$.

Here F and G may be approximated arbitrarily closely in the sup norm by sequences $\{F_n\}$ and $\{G_n\}$ respectively such that G_n and F_n satisfy the hypothesis of the theorem and (2).
We now have

$$\int L(X)dF_n(X) \leqslant \int L(X)dG_n(X).$$

The proof is completed by taking limits.

COMMENT: A number of colleagues have pointed out an alternative proof using integration by parts. However, the proof given here avoids breaking the integral $\int L(X)d(F-G)$ into two pieces.

In addition a referee has pointed out that the inequality in [3], $\lambda(X_0)(X-X_0) \geqslant L(X)-L(X_0)$, where $\lambda(X_0)$ can be shown to be a decreasing function can be used in equation (2) to produce a slightly shorter but a more sophisticated proof. This proof is similar to the proof of case 1 but does not require any approximation.

Finally, it is stated that the open interval in the statement of the theorem can be replaced by any interval. This may be seen by taking appropriate limits of concave functions defined in open sets. The essential point here is that concave functions can be discontinuous at the endpoints of their domain.

References

[1] U.S. Environmental Protection Agency, Air Quality Criteria for Lead, EPA-600/8-77-017, Washington, D.C., December 1977.

[2] Billick, I., Shier, D. and Spiegelman, C., Sensitivity of Observed Blood Lead to Measurement Errors. Unpublished manuscript, 1980.

[3] Loeve, M., *Probability Theory.* Third Edition. Van Nostrand Co., Inc., p. 159, 1963.